Mètode MAREA: Mètode d'Auditoria i Recomanació d'Entorns Analítics

Mireia Sempere Mateu

Introducció

El mètode MAREA (*Method for Analytical Environment Review and Advisory*, Mètode d'auditoria i recomanació d'entorns analítics) és una guia estructurada dissenyada per a empreses i organitzacions que volen avaluar el seu entorn analític. El seu objectiu és determinar si l'entorn implementat és adequat a les necessitats presents i futures de l'organització, o si cal introduir modificacions per garantir un millor ajustament als requisits funcionals, tècnics i estratègics i presentar un *technology stack* que s'adapti a la solució escollida.

El mètode MAREA sorgeix com a resultat del Treball de Final de Grau (TFG) del curs 2024-25 de la Mireia Sempere Mateu, estudiant d'Enginyeria de Dades a l'Escola d'Enginyeria de la Universitat Autònoma de Barcelona. L'objectiu principal d'aquest treball és dur a terme una auditoria i formular una proposta de recomanació d'un entorn analític adequat per a l'empresa Marlex. Al llarg del desenvolupament del projecte, s'ha dissenyat i estructurat aquest mètode amb la voluntat de que pugui ser aplicat de manera generalitzada a altres entorns empresarials o organitzacionals amb necessitats similars d'avaluació i millora del seu ecosistema analític.

Les fases que conformen el mètode MAREA són les següents:

1. **Problem Statement**: Identificació i definició precisa de la necessitat principal del client, basada en la comprensió del problema a resoldre i del context en què es troba l'organització.
2. **Estat de l'Art**: Revisió de les tendències, tecnologies i bones pràctiques actuals en entorns analítics per contextualitzar la presa de decisions.
3. **Elecció de la Metodologia per la Recollida de Requisits i per la Posterior Implementació del Projecte**: Selecció del marc metodològic per al desenvolupament de la solució, tenint en compte els recursos de l'empresa.
4. **Recollida de Requeriments i Anàlisi del Context Tècnic**: Recollida d'informació sobre l'empresa, el seu entorn analític actual i les seves condicions tècniques per orientar el disseny de la futura solució.
5. **Elecció d'Arquitectura**: Disseny i justificació de l'arquitectura més adequada per a l'entorn analític segons els requisits identificats i l'estudi previ de l'Estat de l'Art.
6. **Elecció d'Infraestructura**: Avaluació i selecció de la infraestructura òptima (*on-premise* (local), *cloud* (al núvol) o híbrida) en funció dels requisits, el context de l'organització i l'estudi previ de l'Estat de l'Art.
7. **Proposta Final i Conclusions**: Presentació de la proposta final i conclusions complementàries recollides durant l'auditoria.

L'elecció de l'entorn analític es realitza en dues dimensions diferenciades. En la primera (secció 5), s'estudia l'arquitectura més adequada per al projecte de forma independent, evitant que condicionants tècnics o limitacions pròpies de la infraestructura influeixin en la decisió arquitectònica. Un cop definida l'arquitectura òptima segons els requisits funcionals i analítics, es passa a avaluar les diferents opcions d'infraestructura disponibles que s'ajustin a aquest model i que siguin capaces de satisfer les necessitats identificades (Secció 6). Com a resultat del mètode s'acabarà presentant un *technology stack* adeqüat al cas estudiat juntament amb les recomanacios associades.

1 Problem Statement

El primer pas per a poder començar a treballar en buscar la millor solució és centrar-se en la petició del client final. Cal recollir informació rellevant sobre l'empresa o organització; de quin sector pertanyen, quines dimensions té l'empresa i l'equip tècnic, quin *budget* (pressupost) té, i quin és el problema, és a dir, quina és la **necessitat** principal. D'aquesta manera ja es tindrà informació per a seguir amb els passos *Estat de l'Art* (2) i *Elecció de la Metodologia* (3).

2 Estat de l'Art

Per tal de poder formular una recomanació fonamentada, és imprescindible dur a terme un treball previ d'anàlisi. Concretament, cal estudiar l'Estat de l'Art del projecte, identificant iniciatives similars, tendències actuals en el sector i pràctiques adoptades per les empreses competidores. Així mateix, és important estudiar els estàndards i les noves tendències d'**arquitectures** i opcions de models i serveis d'**infraestructura**, tenint en compte l'evolució tecnològica vinculada a l'àmbit específic del projecte. Només tenint aquesta visió global i contextualitzada es pot proposar una recomanació que sigui realment actualitzada, rellevant i adequada a les necessitats específiques del cas analitzat.

A Juny del 2025, data del mètode MAREA, en l'Estat de l'Art s'ha estudiat que les principals arquitectures que s'han tornat més populars i reconegudes al llarg de la història:

- **Inmon - *Data Warehouse*** (*top-down*): proposa construir primer un DW centralitzat en 3NF (normalitzat) anomenat *Enterprise Data Warehouse*, del qual se'n deriven els *Data Marts*. És un enfocament robust i escalable [2] enfocat a la màxima integració però el seu desenvolupament és lent degut a la normalització i a la planificació prèvia.
- **Kimball - *Data Warehouse*** (*bottom-up*): parteix de Data Marts dimensionals (amb *star* o *snowflake schemas*) que s'integren posteriorment en un DW corporatiu. Permet obtenir valor ràpidament [2] ja que l'objectiu és la usabilitat i es prioritza un enfocament departamental amb dades des-normalitzades per facilitar l'anàlisi. [6]
- **Arquitectura Mixta (Inmon + Kimball) - *Data Warehouse***: combina els avantatges dels enfocaments d'Inmon i Kimball per adaptar-se millor a les necessitats d'entorns empresarials complexos. Es parteix d'un DW corporatiu centralitzat i normalitzat (seguint Inmon), però també es desenvolupen *Data Marts* departamentals amb esquemes dimensionals (seguint Kimball) per oferir agilitat analítica, amb una capa de dimensions compartides entre *Data Marts*. Aquesta estratègia híbrida permet una alta integració i governança de les dades i alhora manté l'accessibilitat i usabilitat analítica [2].
- ***Data Vault:*** plantejat per Dan Linstedt [10, 11], és una arquitectura modular i flexible pensada per a entorns amb grans volums de dades i canvis freqüents. Organitza la informació en tres tipus de taules: Hubs (claus de negoci), Links (relacions entre claus) i Satellites (atributs i historials), facilitant la traçabilitat i l'escalabilitat.
- ***Data Lake:*** emmagatzema dades en format cru, semi-estructurat o no estructurat, permetent l'ingestió massiva per a casos de *Big Data*, anàlisi exploratòria i *Machine Learning*. Tot i així, sense una bona governança i catalogació, pot tornar-se un *data swamp* (pantà de dades).
- ***Data Lakehouse:*** aquesta arquitectura permet emmagatzemar dades tant estructurades com no estructurades en un únic sistema, aprofitant els avantatges del *Data Lake* (*schema-on-read*, cost reduït, flexibilitat, ingestió massiva) i del *Data Warehouse* (*schema-on-write*, esquema estructurat, alta governança, qualitat de dades). És especialment útil en entorns híbrids que combinen necessitats de *Business Intelligence* (BI) i *Machine Learning* avançat, i permet fer anàlisi en temps casi real o en *batch* sense duplicació de dades.

Aquestes arquitectures seran analitzades en la secció *Elecció d'Arquitectura*, punt 5.

3 Elecció de la Metodologia per la Recollida de Requisits i per la Posterior Implementació del Projecte

Per a auditorar adecuadament és fonamental definir des d'un inici una metodologia que permeti un correcte desenvolupament del treball. Aquesta elecció ha de tenir en compte tant la naturalesa del projecte com els recursos disponibles, la necessitat de flexibilitat, la col·laboració amb l'empresa i la planificació temporal. També és fonamental tenir en compte les tendències metodològiques observades en projectes similars, per tal de fonamentar la selecció en criteris actuals i contrastats. Tota aquesta informació es recull en el primer i segon pas *Problem statement* i *Estat de l'Art*. Per això en aquest punt s'ha d'escollir una metodología per la recollida de requisits i una altra per la futura implementació del projecte, també molt lligada al context i els recursos de l'empresa.

Un exemple d'opcions de metodologies són les següents:

- **Recollida de requisits**:
  - **Kimball**: Metodologia *bottom-up* [2] orientada a la creació de *Data Marts* dimensionals a partir dels processos de negoci més rellevants. La recollida de requisits se centra en entendre les preguntes clau del negoci, indicadors i mètriques que els usuaris necessiten analitzar, i per tant, es centra des d'un principi en els requisits departamentals. A partir d'això, durant el desenvolupament es fa la unió entre els requisits de cada *Data Mart* i s'extrapolen a nivell d'empresa per a construir l'arquitectura de bus (*Bus architecture* [2]) que connecta tots els departaments.
  - **Inmon**: Metodologia *top-down* per a la construcció d'un *Data Warehouse* corporatiu, centralitzat i altament normalitzat. La recollida de requisits comença amb una anàlisi global i exhaustiva de tota l'organització, amb

una visió transversal orientada a la integració completa de les dades de l'empresa. Posteriorment, s'aprofundeix en les necessitats particulars de cada departament mitjançant la derivació de *Data Marts* a partir de la base corporativa comuna [2].

En aquells contextos on es prioritza entendre i cobrir ràpidament les necessitats analítiques específiques de cada departament i obtenir resultats funcionals des de les primeres fases, el model de Kimball és més adequat, ja que orienta la recollida de requisits cap als processos de negoci concrets i permet definir els *Data Marts* de forma incremental. En canvi, si l'objectiu és construir una visió corporativa global i plenament integrada des del primer moment, amb un fort control estructural sobre la qualitat i la normalització de les dades, aleshores l'enfocament d'Inmon és preferible, ja que requereix una recollida inicial de requisits molt més extensa i transversal.

- **Implementació del projecte (desenvolupament i migració)**:
  - **Waterfall** (model en cascada): Metodologia tradicional i seqüencial en què el desenvolupament es divideix en fases clares: anàlisi, disseny, implementació, proves i manteniment. Cada fase es completa abans de passar a la següent. És adequada per a projectes amb requisits estables i molt ben definits però no es té en compte que els requeriments poden canviar al llarg del projecte [12].
  - **Scrum**: Metodologia àgil [7] basada en iteracions anomenades *sprints*, en què es planifica, desenvolupa i entrega una part funcional del producte. El treball es gestiona a través d'un *Product Backlog*, una llista prioritzada de funcionalitats definida pel *Product Owner*, de la qual se'n seleccionen els ítems que es desenvolupen en cada sprint. Scrum promou la col·laboració constant amb l'usuari, reunions de seguiment regulars i una alta capacitat d'adaptació als canvis. És especialment adequada per a projectes amb requisits canviants i entorns iteratius però requereix un equip d'almenys cinc membres per a cobrir tots els rols fonamentals [9].
  - **Kanban**: Metodologia àgil [7] visual que utilitza un tauler per representar l'estat de cada tasca i controlar el flux de treball. A diferència de Scrum, no es divideix en iteracions tancades, sinó que permet una gestió contínua i més flexible. A més a més, no requereix d'un nombre mínim de membres perquè no defineix rols ni estructures fixes. Posa l'èmfasi en la limitació del treball en curs (*Work In Progress*, WIP) i és molt útil per a equips amb demandes canviants i processos més estables o recurrents [9].

La selecció de la metodologia de desenvolupament ha d'adaptar-se a les característiques del projecte, l'equip disponible i el grau de flexibilitat temporal i funcional que es pugui assumir. El model Waterfall, per exemple, tot i ser estructurat i fàcil d'implementar, només es recomana en projectes petits, amb una baixa complexitat tècnica i quan no es preveguin canvis significatius en els requisits definits al principi.

En canvi, en projectes de major envergadura, amb equips complets i terminis de lliurament marcats, és preferible aplicar Scrum, ja que estructura el desenvolupament en iteracions curtes que permeten ajustar i prioritzar funcionalitats progressivament, afavorint la transparència i la col·laboració contínua amb el client.

Finalment, en aquells projectes amb més flexibilitat temporal i/o amb equips més reduïts, on es preveuen canvis freqüents en les funcionalitats o no és possible establir *sprints* tancats, la metodologia Kanban és una opció òptima. La seva capacitat per gestionar fluxos de treball continus, sense iteracions predefinides ni dependència d'una estructura rígida, la fa especialment útil en entorns canviants o amb tasques recurrents.

En funció de tots aquests factors s'ha d'escollir la metodologia ideal per a que la recollida de requisits i el desenvolupament es duguin a terme d'acord amb el context i els recursos de l'oorganització.

4 Recollida de Requeriments i Anàlisi del Context Tècnic

A partir de la metodologia escollida per a la recollida de requisits, s'organitzen les reunions necessàries amb el personal de l'empresa i els *stakeholders* per identificar i documentar totes les necessitats del projecte. Cal recollir requisits tant funcionals com tècnics:

- **Requisits del producte final**: informació que es vol visualitzar, departaments implicats, objectius de negoci i ús previst.
- **Requisits tècnics** (especialment en entorns analítics):
  1. Infraestructura on està allotjat l'entorn analític (màquines, cost, localització).
  2. Volum actual de les bases de dades i previsió de creixement a 3 anys vista.
  3. Fonts de dades actuals i orígens previstos en el futur.
  4. Tipologia de Business Intelligence: analítica convencional, *machine learning* clàssic o avançat i quines eines utilitzen actualment.
  5. Tipus de dades i de còmput: càrregues periòdiques o fluxos *near real-time*.
  6. Tipus de càlcul requerit: puntual (models), periòdic (ETLs, informes), o en temps de feina (consultes *ad hoc*).
  7. Nombre de membres de l'equip tècnic i plans d'inversió en recursos a 3 anys vista (noves incorporacions o contractació de recursos externs).

8. Formació i especialització de l'equip: eines, tecnologies i rols dominats.
9. Identificació de colls d'ampolla tècnics i preocupacions principals.

5 Elecció d'Arquitectura

A partir de les opcions arquitectòniques analitzades a l'Estat de l'Art, es procedeix a seleccionar l'arquitectura més adequada tenint en compte tota la informació recollida durant la fase d'obtenció de requisits. Aquesta fase té dues parts diferenciades; en primer lloc, s'ha elaborat un **arbre de decisió d'arquitectura** (**ADT**, *Architecture Decision Tree*) que permet determinar quines arquitectures són viables en funció de la situació actual de l'empresa i els recursos tècnics disponibles. En segon lloc, s'ha definit una **matriu de ponderació de criteris arquitectònics** (**ACWM**, *Architectural Criteria Weighting Matrix*) que s'explicarà més endavant.

L'ADT (Figura 1) es fonamenta principalment en quatre criteris tècnics clau obtinguts a la recollida de requeriments: el *Volum de la base de dades* (punt 2), la *tipologia de Business Intelligence* (punt 4), el *Tipus de dades* (punt 5) i el *nombre de membres de l'equip tècnic o la predisposició de l'empresa d'invertir en recursos* (punt 7).

Aquest arbre de decisió permet valorar la opció arquitectònica que s'adapti a les necessitats d'analítica però tenint en compte les limitacions de l'empresa com pot ser els recursos o el temps de desenvolupament. L'arbre de decisió següent ha estat actualitzat al juny de 2025, data de referència del present projecte, amb les opcions d'arquitectura estudiades a l'Estat de l'Art (punt 2):

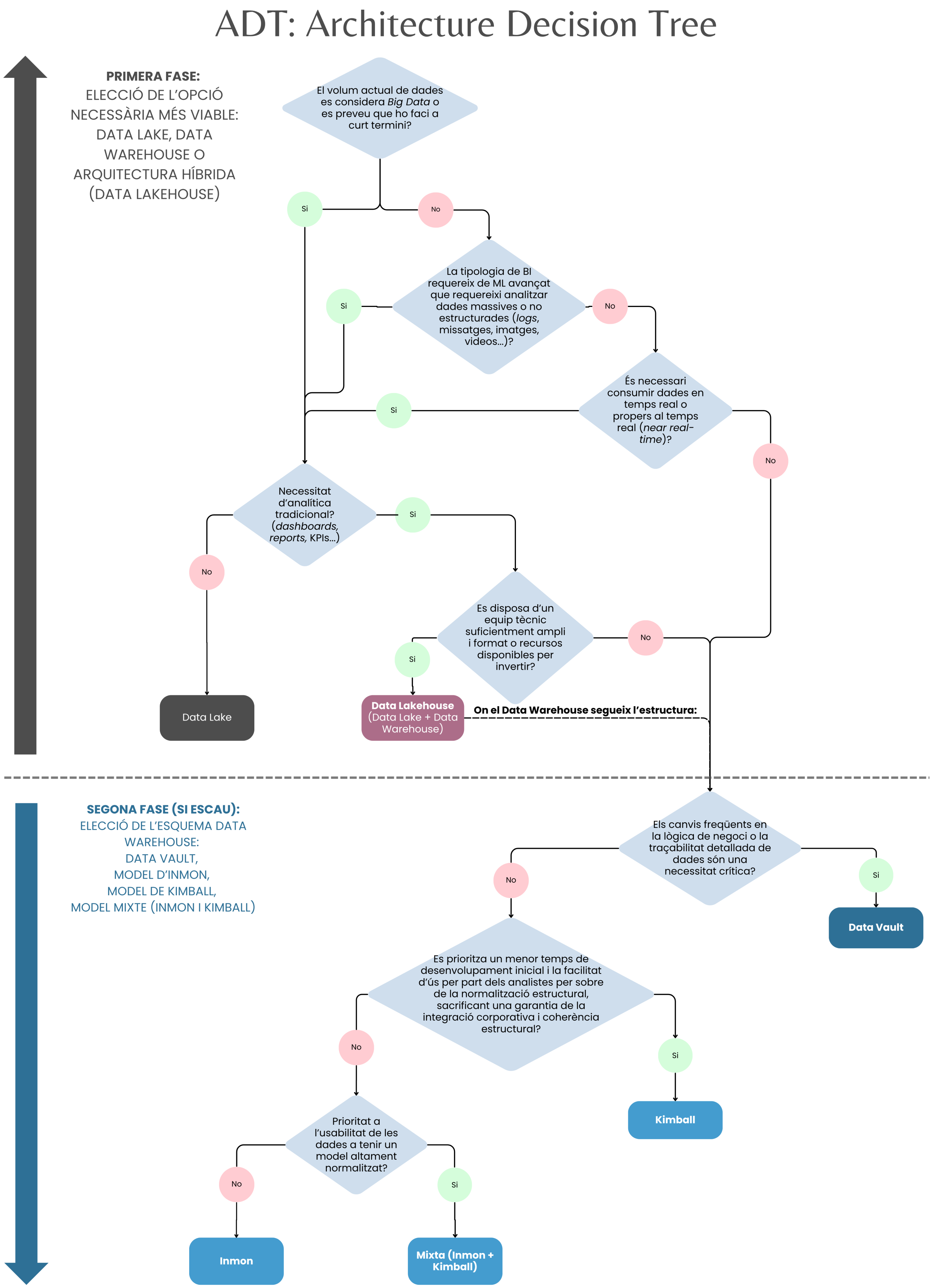

Fig. 1: Arbre de Decisió d'Arquitectura (ADT) per estudiar la opció més viable.

Un cop obtinguda una primera proposta mitjançant aquest ADT, es complementa amb l'anàlisi de la matriu de ponderació de criteris arquitectònics (ACWM, *Architectural Criteria Weighting Matrix*). Aquesta matriu permet comparar de manera estructurada i objectiva les diverses alternatives a partir d'un conjunt de criteris arquitectònics rellevants. Les limitacions de l'empresa com els recursos o el temps de desenvolupament no es tenen en compte, i per tant l'objectiu d'aquesta ACWM és valorar l'arquitecura ideal en funció de les necessitats de l'empresa o organització. Els criteris que s'han definit són els següents:

1. Facilitat d'ús per part dels usuaris de BI
2. Integració de fonts múltiples
3. Gestió d'històrics
4. Escalabilitat i flexibilitat
5. Facilitat de manteniment
6. Resiliència davant canvis de negoci
7. Suport a dades atòmiques i agregades
8. Capacitat per treballar amb dades recents o en temps real
9. Simplicitat de l'arquitectura
10. Governança i seguretat

Cada arquitectura és avaluada en una escala de 0 a 11 (0-1-4-7-11) en cadascun d'aquests criteris. Aquesta escala permet quantificar el grau de cobertura funcional que cada solució ofereix respecte a cada criteri, segons el següent significat:

- 0: No ofereix cap tipus de cobertura
- 1: Cobertura molt limitada
- 4: Cobertura funcional bàsica
- 7: Cobertura funcional avançada
- 11: Cobertura excel·lent

A més, amb l'objectiu de reflectir el pes específic que pot tenir cada criteri en funció de la seva importància relativa dins del context d'anàlisi, és possible aplicar una ponderació individual (w_i) per a cadascun d'ells. Alternativament, si es considera que tots els criteris tenen la mateixa rellevància, es poden assignar pesos uniformes (valor 1) a tots.

La puntuació final assignada a cada arquitectura s'obté mitjançant la fórmula següent:

$$\text{Puntuació final de cada opció} = \sum_{i=1}^{n} (w_i \cdot s_i)$$

on:

- w_i és el pes (ponderació) assignat al criteri i
- s_i és la puntuació obtinguda per l'opció en el criteri i
- n és el nombre total de criteris considerats en l'anàlisi

A continuació, es mostra un exemple de la matriu ACWM, en què s'ha considerat que tots els criteris tenen la mateixa importància relativa. En conseqüència, s'ha assignat un pes unitari (1) a cadascun d'ells.

Criteri	**Pes**	**Kimball**	**Inmon** (Hub&Spoke)	**Mixta** (Inmon+Kimball)	**Data Vault**	**Lakehouse**
Facilitat d'ús per usuaris BI	1	11	4	11	1	7
Integració de fonts múltiples	1	4	11	11	11	11
Gestió d'històrics	1	4	11	11	11	11
Escalabilitat i flexibilitat	1	4	4	7	11	11
Facilitat de manteniment	1	7	4	7	4	11
Resiliència davant canvis de negoci	1	1	7	11	11	11
Suport a dades atòmiques i agregades	1	11	7	11	11	11
Temps real o dades recents	1	4	1	4	1	11
Simplicitat de l'arquitectura	1	11	4	7	1	4
Governança i seguretat	1	11	7	11	4	7
Total (sobre 110)		**68**	**60**	**91**	**66**	**95**

TAULA 1: EXEMPLE D'ACWM, *Architectural Criteria Weighting Matrix*.

En aquest exemple es pot observar quines arquitectures obtenen una millor puntuació en funció dels criteris establerts. Aquesta matriu és extensible i pot adaptar-se a les necessitats específiques del projecte, de manera que es poden incorporar nous criteris a partir de l'anàlisi prèvia i dels objectius definits. Així mateix, és possible ajustar els pesos dels criteris per posar l'èmfasi en aquells aspectes considerats més rellevants, permetent així explorar diferents escenaris i tendències.

5.1 Recomanació d'arquitectura: *technology stack*

A partir dels resultats obtinguts en la matriu ACWM i de l'arbre de decisió arquitectònica (ADT), s'identifica tant l'arquitectura més viable com l'arquitectura ideal. En cas que aquestes divergeixin, es duu a terme una anàlisi crítica per determinar quins requisits, limitacions tècniques o condicions contextuals impedeixen implementar l'alternativa òptima identificada a la matriu. Aquest enfocament permet seleccionar una arquitectura que no només sigui la millor des d'un punt de vista tècnic, sinó que també sigui realista i viable d'implementar en l'entorn actual del projecte.

A més a més, a partir de les capes definides de l'arquitectura escollida, també cal plantejar els processos ETL (*Extract, Transform and Load*) o ELT (*Extract, Load and Transform*) per proporcionar una recomanació integral.

A partir de l'elecció escollida s'elabora un *technology stack* gràfic i visual per representar tot l'esquema arquitectònic definit. Un exemple de *tecknology stack* d'una arquitectura *Data Lakehouse* és el següent:

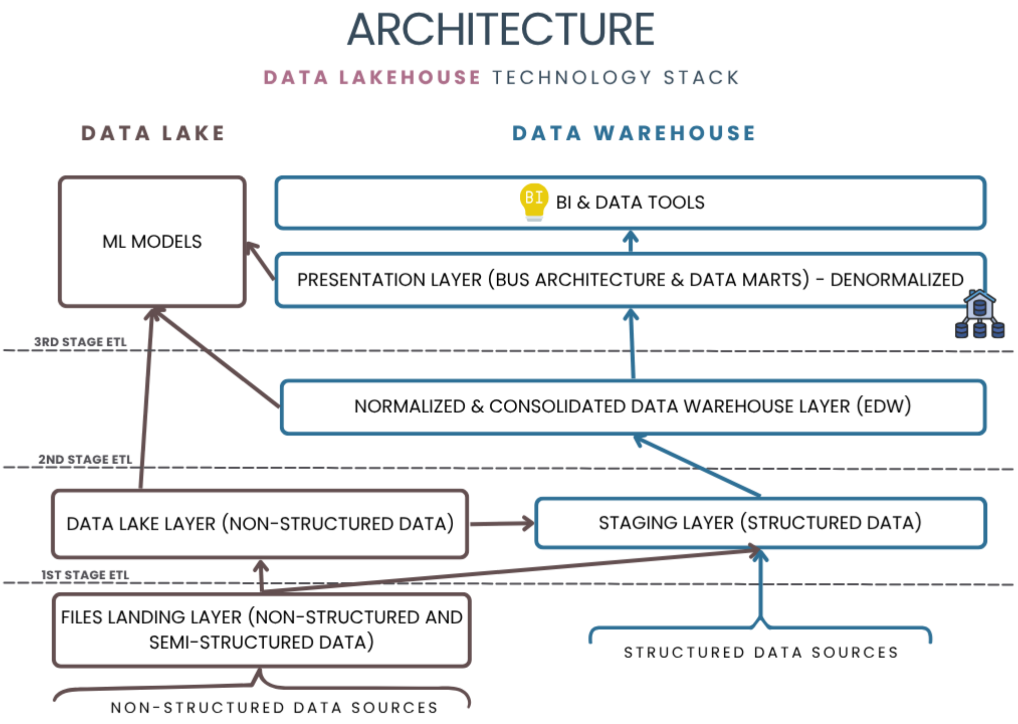

Fig. 2: *Technology Stack*: *Data Lakehouse*

6 ELECCIÓ D'INFRAESTRUCTURA

A partir de l'estudi previ de l'Estat de l'Art, de les opcions disponibles al mercat i de les tendències actuals en entorns analítics, es procedeix a la selecció de la infraestructura més adequada per allotjar la solució proposada. Aquesta decisió ha d'estar guiada tant per aspectes econòmics com per criteris tècnics i organitzatius.

En primer lloc, cal valorar el context actual des d'on parteix l'empresa o organització. Per a fer-ho, es fa un anàlisi general de les característiques de l'entorn i es calcula el **cost total de propietat (TCO)** [13] de l'entorn actual a tres anys vista. Aquest càlcul es fa estimant els costos CapEx (*Capital Expenditure*) i OpEx (*Operational Expenditure*) a tres anys; despeses de maquinari, llicències, manteniment, personal, cost de desplegament i migració, etc. Si la infraestructura analítica actual comparteix recursos amb entorns operacionals, cal avaluar l'impacte i el cost de separar aquests entorns.

Un cop feta aquesta anàlisi de l'entorn actual (i, si escau, de separació d'entorns), es passa a valorar altres alternatives d'infraestructura tenint en compte el tipus d'aqitectura escollit, seguint el mateix procés d'analitzar les caracteristiques de cada opció i els respectius TCOs. Les opcions d'infraestructures valorades a data de la creació del mètode són:

- ***On-premise* dedicat**

- ***Cloud* privat no gestionat**
- ***Cloud* privat gestionat**
- ***Cloud* públic**, comparant entre els tres principals proveïdors del moment en el que es dugui a terme l'auditoria

En el cas del *cloud públic*, és fonamental considerar els diversos models de servei disponibles:

- **IaaS** (Infraestructura com a servei)
- **PaaS** (Plataforma com a servei)
- **SaaS** (Software com a servei)

Com s'ha mencionat, per garantir una comparació equitativa, és fonamental que totes les solucions i serveis proposats es considerin dins del context d'un **technology stack complet** (és a dir, incloent components de comput, emmagatzematge, integració, governança, visualització, etc.), aprofitant que cada opció de *technology stack* (*on-premise*, al *cloud*, etc) cobreixi tota l'elecció d'arquitectura definida al punt anterior. Així es garanteix que el càlcul del TCO sigui coherent entre diferents opcions.

A més del TCO, es valoren diversos criteris tècnics i operatius mitjançant una matriu de ponderació de criteris d'infraestructura (ICWM, *Infrastructure Criteria Weighting Matrix*), que permet ponderar cada infraestructura de manera estructurada. Els criteris considerats són:

1. Governança i seguretat
2. TCO (cost total de propietat)
3. Coneixement intern disponible
4. Dependència del proveïdor (a menys escala, més dependència)
5. Flexibilitat i personalització
6. Disponibilitat d'eines integrades per a *machine learning*
7. Flexibilitat tècnica per a *machine learning* personalitzat
8. Integració amb l'eina de visualització i analítica que s'utilitzi a l'organització
9. Facilitat de migració i instal·lació
10. Escalabilitat
11. Compatibilitat amb ETLs existents
12. Temps de migració i desplegament
13. Temps d'aprenentatge

Cada opció d'infraestructura s'avalua segons aquests criteris, utilitzant el mateix sistema de puntuació que la matriu de ponderació de criteris d'arquitectura ACWM. En aquest punt es recomana ponderar els criteris segons prioritat degut a que depèn molt de l'objectiu de l'entorn analític. També es pot considerar afegir mes criteris.

Un exemple d'aplicació pràctica és la matriu següent 2, amb dos escales de pesos diferents. En aquesta matriu només es mostren les opcions *cloud*. Els totals resultants de cada opció seguint la fòrmula estan resumits a la taula 3:

Criteri	**Pes** (1-5), (0-11)	**On-Premise** Entorn Actual	**Azure** Fabric (SaaS)	**Azure** Synapse (PaaS)	**Azure** SqlDB +MI (PaaS)	**Azure** VMs + SQL (IaaS)	**AWS** Redshift (SaaS)	**AWS** RDS +SQL (PaaS)	**AWS** EC2 +SQL (IaaS)	**Google** BQuery (SaaS)	**Google** Cloud SQL (PaaS)	**Google** GCE +SQL (IaaS)
Coneixement intern disponible	3, 4	11	4	4	4	7	0	4	4	0	1	1
Dependència del proveïdor (més petit, més depenent)	3, 4	7	1	4	4	7	1	4	7	0	4	7
Flexibilitat i personalització	3, 4	11	1	4	4	11	1	4	11	1	4	11
Eines integrades per a ML	3, 4	0	11	11	4	1	4	1	1	7	4	1
Flex. tècnica per a ML personalitzat	3, 4	4	1	7	4	11	1	4	11	4	4	11
Integració amb amb l'eina de visualització	4, 7	11	11	11	11	7	4	4	4	7	4	4
Facilitat de migració i instal·lació	4, 7	4	4	7	7	11	0	4	7	0	4	7
Escalabilitat	4, 7	4	7	11	7	4	11	7	4	11	7	4
Compatibilitat amb ETL existents	4, 7	11	1	11	7	11	1	4	7	0	4	7
Temps de migració i desplegament	4, 7	1	11	7	7	1	11	7	4	7	7	4
Temps d'aprenentatge	4, 7	7	1	4	7	11	1	4	7	1	4	7
TCO	5, 11	11	11	0	7	4	11	4	1	0	11	0
Governança i seguretat	5, 11	4	7	7	7	4	4	4	4	4	4	4

TAULA 2: MATRIU ICWM EXEMPLE DESGLOSSADA PER CRITERIS I PONDERACIÓ DEL PES

Tipus d'avaluació	On-Premise Entorn Actual	Azure Fabric (SaaS)	Azure Synapse (PaaS)	Azure SQL DB/MI (PaaS)	Azure VMs + SQL (IaaS)	AWS Redshift (SaaS)	AWS RDS + SQL (PaaS)	AWS EC2 + SQL (IaaS)	Google BQuery (SaaS)	Google Cloud SQL (PaaS)	Google GCE + SQL (IaaS)
Sense ponderar	86	71	88	80	90	50	55	72	42	62	68
Ponderat (1–5)	326	284	329	314	331	208	211	259	160	246	245
Ponderat (0–11)	563	515	554	556	551	389	366	422	274	443	399

TAULA 3: RESULTAT DE LA MATRIU EXEMPLE ICWM (TAULA 2)

.

A partir dels resultats de la matriu es poden veure els resultats i la tendència al tenir en compte els pesos o augmentar l'escala d'aquests pesos. Juntament amb els millors resultats i valorant les prioritats de l'empresa (com per exemple podria ser el TCO), s'assegura una elecció equilibrada i ajustada a la seva realitat.

6.1 Recomanació final: *technology stack*

A partir de la opció escollida, s'ha de presentar el *technology stack* indicant la infraestructura a on estarà allotjat i si és un servei *cloud*, el tipus de servei.

Un exemple de *tehnology stack* d'Azure amb un servei *Infrastructure As A Service* és:

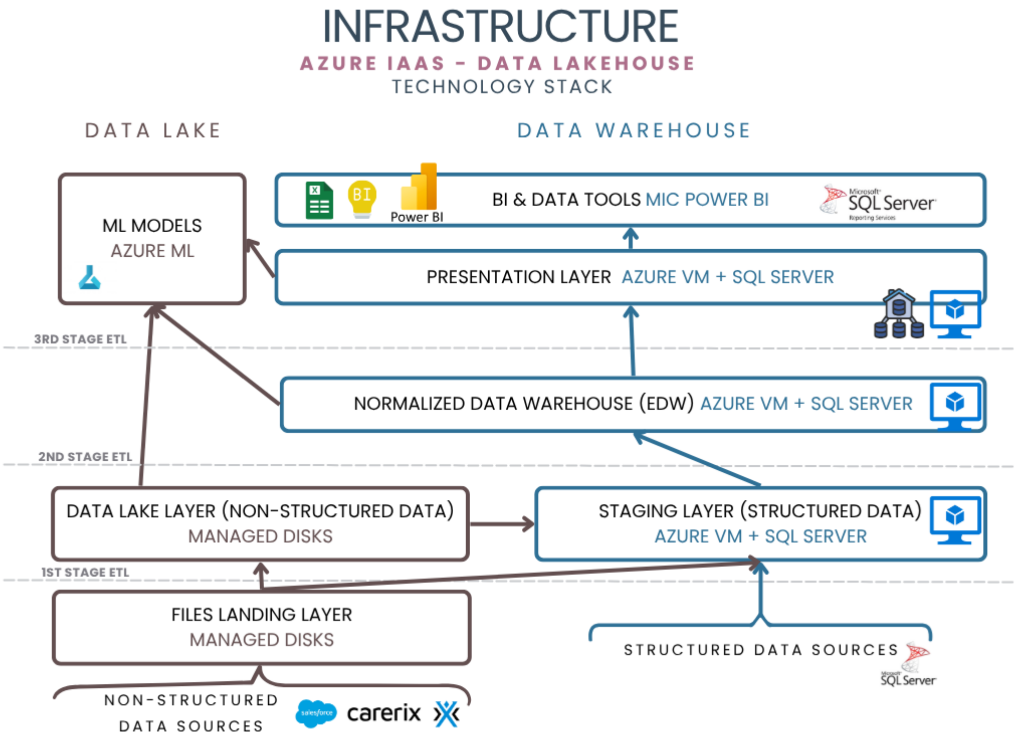

Fig. 3: *Technology Stack*: *Data Lakehouse* allotjat a Azure IaaS.

7 Proposta Final i Conclusions

En aquesta secció es presenten els dos *technology stacks* (o un que integri els dos); el que detalla l'arquitectura i el que detalla la infraestructura més adients per a l'organització, acompanyada de les justificacions tècniques corresponents. A més, es poden incloure recomanacions addicionals i conclusions rellevants derivades de les observacions recollides durant el procés d'auditoria.

Referències

[1] Modular.Click, "Marlex Great People — Selección de Personal - RRHH." [Online]. Available: https://www.marlex.net/

[2] R. Kimball and M. Ross, The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling, 3rd ed. Hoboken, NJ: Wiley, 2013. [Online]. Available: http://cds.cern.ch/record/1623624

[3] D. Linstedt and M. Olschimke, Building a Scalable Data Warehouse with Data Vault 2.0. Waltham, MA: Morgan Kaufmann, 2015. [Online]. Available: https://dl.acm.org/citation.cfm?id=2901170

[4] S. Hoberman, M. Blaha, B. Inmon, G. Simsion, and C. Lehn, Data Modeling Made Simple: A Practical Guide for Business and IT Professionals, 2nd ed. Technics Publications, 2009. [Online]. Available: https://www.amazon.com/Data-Modeling-Made-Simple-Professionals/dp/0977140067

[5] Scrum.org, "What is Scrum?" [Online]. Available: https://www.scrum.org/learning-series/what-is-scrum/

[6] M. Ross, "The 10 Essential Rules of Dimensional Modeling," Kimball Group, Jan. 26, 2016. [Online]. Available: https://www.kimballgroup.com/2009/05/the-10-essential-rules-of-dimensional-modeling/

[7] R. de Tena, "Metodología AGILE: qué es y en qué consiste," Holded Blog, Apr. 12, 2022. [Online]. Available: https://www.holded.com/es/blog/metodologia-agile

[8] G. E. Silva Peñafiel, V. M. Zapata Yánez, K. P. Morales Guamán, and L. M. Toaquiza Padilla, "Análisis de metodologías para desarrollar Data Warehouse aplicado a la toma de decisiones," Ciencia Digital, vol. 3, no. 3.4, pp. 397–418, 2019. [Online]. Available: https://doi.org/10.33262/cienciadigital.v3i3.4.922

[9] Development That Pays. “Scrum vs Kanban - What’s the Difference?” *YouTube*, Jan. 18, 2017. [Online]. Available: https://www.youtube.com/watch?v=rIaz-l1Kf8w

[10] Linstedt, D., Graziano, K., & Hultgren, H. *The Business of Data Vault Modeling*. [Online]. Available: https://www.amazon.es/Business-Data-Vault-Modeling/dp/143571914X

[11] Linstedt, D., & Graziano, K. *Super Charge Your Data Warehouse: Invaluable Data Modeling Rules to Implement Your Data Vault: Volume 1*. [Online]. Available: https://www.amazon.es/Super-Charge-Your-Data-Warehouse/dp/1463778686

[12] B. Atlassian, “Waterfall Methodology: A Comprehensive Guide,” *Atlassian*, Apr. 24, 2025. [Online]. Available: https://www.atlassian.com//project-management/waterfall-methodology

[13] Wikipedia contributors, “Coste total de propiedad,” *Wikipedia, la Enciclopedia Libre*, Jan. 20, 2025. [Online]. Available: https://es.wikipedia.org/wiki/Coste_total_de_propiedad